Approaching descriptive and theoretical truth

Kuipers, Theo A.F.

Published in:
Erkenntnis, vol. 18 (3), 343-378

Publication date:
1982

THEO A. F. KUIPERS

APPROACHING DESCRIPTIVE AND THEORETICAL TRUTH*

1. INTRODUCTION

According to Graham Oddie's recent review (1981) of the literature on verisimilitude, or truthlikeness, there are essentially three different new approaches to the problem of verisimilitude after the breakdown, in 1974, of Popper's original definition. All three projects happen to be co-productions: Miller-Popper, Niiniluoto-Tuomela and Tichý-Oddie. The first-mentioned authors came with the first ideas whereas the second-mentioned extended and amended on them.

In this paper we will argue that in the whole discussion on verisimilitude up till now, including Oddie's review, there have been conflated two different questions, the conflation arising from the generally held assumption that 'the truth' is to be identified with the true description of the actual world.

By making the distinction between true descriptions and true theories we will make clear that there are essentially two problems of verisimilitude, viz. that of descriptive and that of theoretical verisimilitude. It will be argued that Niiniluoto has given the adequate definition of descriptive verisimilitude, i.e. of judgements of the form 'description B is closer to the

Erkenntnis **18** (1982) 343–378. 0165–0106/82/0183–0343 $03.60

true description than description A'. Moreover, we will see that Miller's definition, provided it is not applied to the true description (as Miller does) but to the true theory, in which form it was independently found by the present author, gives the adequate solution to the problem of theoretical verisimilitude, i.e. it is the basis for judgements of the form 'theory B is closer to the true theory than theory A'.

Our account will be restricted to so-called *qualitative* (or deductive) judgments of verisimilitude and to *propositional* languages, because all main points can already be made within the context of these restrictions.

Starting from a suggestive example we will introduce in Section 2 the descriptive and the theoretical point of view, together with the corresponding truth-notions. These two points of view give rise in a natural way to two different problems of verisimilitude and they suggest also a clear distinction between descriptive and theoretical statements. Sections 3 and 4 present the solutions to the problems of theoretical and descriptive verisimilitude, respectively. In Section 5, and three appendices, the two solutions, especially the 'theoretical' one, are compared with those known in the literature. The introduced truth-notions and the distinction between theoretical and descriptive statements happen to shed, unintended and unexpected, new light on other traditional problems as well. As an example of this the subject of explanation and prediction is treated in Section 6.

In Section 7 we will present and evaluate some methodological rules in the light of the aim of approaching the theoretical truth. It will turn out that Lakatos' rule of falsification, except for its 'novelty-requirement', corresponds to the justifiable, so-called Rule of Success. It will also be shown that this rule needs to be supplemented by a Rule of Content (or Strength), except in the (frequently occurring) cases where the relevant theories are of a particular kind, here called partition-theories. In Section 8 we conclude with some final remarks.

The underlying point of departure of the whole article is the (naive) structuralist view of theories. It is supplemented with statement-formulations and truth-notions, that turn out to imply, as a rule, explications of much informal, especially Popperian, terminology and intuitions.

2. THE DESCRIPTIVE AND THE THEORETICAL POINT OF VIEW

Consider the following electric circuit:

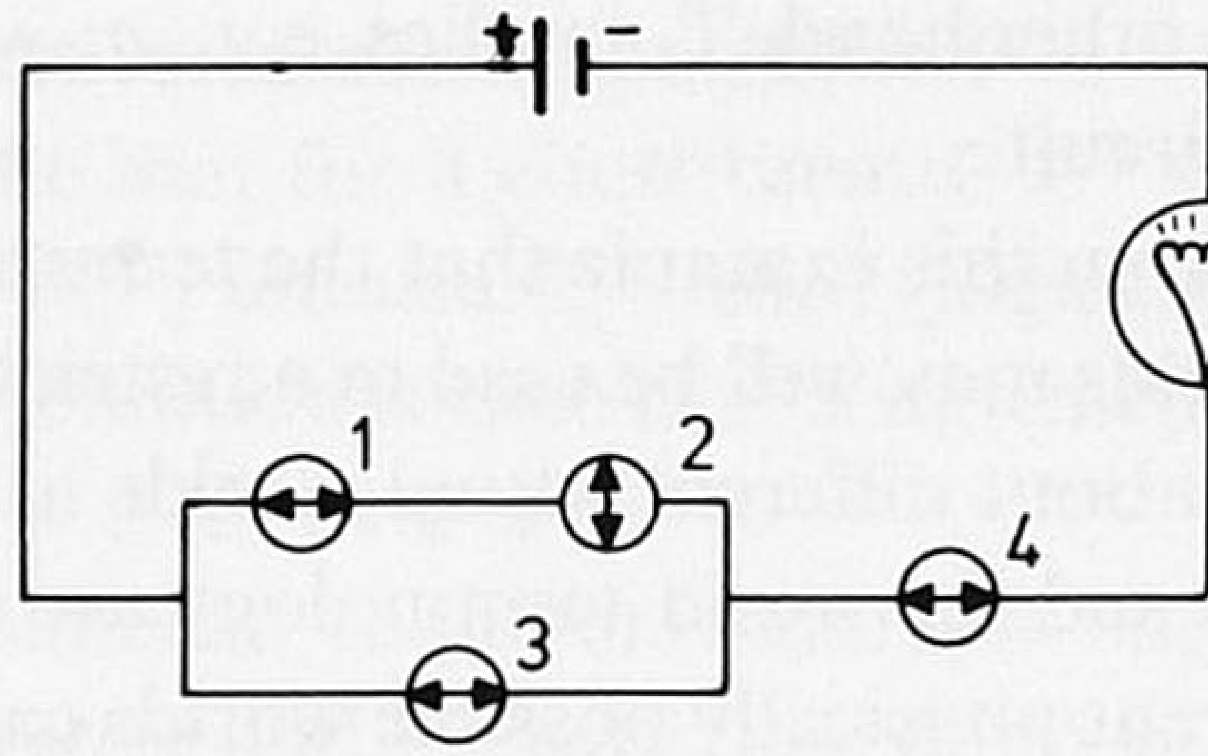

Fig. 1.

Let p_i indicate that switch i is on ($\leftrightarrow$) and $-p_i$ that it is off ($\updownarrow$). Let q ($-q$) indicate that the bulb lights (does not light).

Assuming that the bulb lights indeed in the situation depicted in Figure 1, *the true description* of the represented state of the circuit, or *the actual world*, is obviously given by the statement ('c' refers to the circuit)

(D_c) $\quad p_1 \mathbin{\&} -p_2 \mathbin{\&} p_3 \mathbin{\&} p_4 \mathbin{\&} q$

On the other hand, assuming that the bulb is not defect and that there is enough voltage, *the true theory* about the possible states of the circuit, or *the physically possible worlds*, is obviously given by the statement

(T_c) $\quad (((p_1 \mathbin{\&} p_2) \vee p_3) \mathbin{\&} p_4) \leftrightarrow q$

Our easy way of talking about *the* true description and *the* true theory suggests that there are truth-notions such that any other statement is a false description/a false theory, if it is a description/a theory at all.

But there must be more truth-notions for we would not hesitate to call e.g. $p_1 \mathbin{\&} -p_2$ a true statement and $p_1 \mathbin{\&} p_2$ a false statement. Moreover, we also would not hesitate to call e.g. $q \rightarrow p_4$ a true law and $q \rightarrow p_1$ a false law about the circuit.

However, we would hesitate to call the first couple of statements ($p_1 \mathbin{\&} -p_2$, $p_1 \mathbin{\&} p_2$) laws, let alone true laws. Hence, in addition to different truth-notions, there seems to be also some particular kind of statement which may be true or false laws (and theories). In this connection it is important to note that T_c is incomplete (i.e. not implying each statement or its negation) and, in particular, does not imply p_i or $-p_i$ for any i, nor q or $-q$. Hence, it does not imply conjunctions of such statements, let alone the true description D_c. On the other hand, T_c implies, e.g., $q \rightarrow p_4$, which we called a true law about the circuit.

We may also learn from this example that the terminology of actual and physically possible worlds may well be used in a restricted context and that it makes sense to talk about different actual worlds at different moments. However, the example and the world-terminology are misleading as far as they suggest that different physically possible worlds cannot be realized, by nature or experimentator, at the same time. Another example may help to eliminate this association of the world-terminology. Although there will exist at a certain moment more than one pendulum, not all physically possible pendula will exist. Nevertheless, the true theory about pendula is about all physically possible ones. On the other hand, the true description will be related to one of the existing pendula, 'the actual one', i.e. the one which is actually described.

From the above exposition it will already be clear that our point of departure is essentially the naive structuralist view of theories (developed by Suppes, Sneed and Balzer, and Stegmüller (1979)), i.e. the structuralist view without the (theory-relative) distinction between theoretical and non-theoretical terms. The structuralist possible models correspond to the logically possible worlds, an intended application corresponds to an actual world and because structuralist models are assumed to characterize the intensionally specified intended applications they pretend to capture the physically possible worlds.

Going back to the circuit-example we can now say that it suggests us how to supplement this naive structuralist view with corresponding statement-formulations and relevant truth-notions, at least as far as (scientific) contexts are concerned that can be treated within an (interpreted) propositional language. We start with some more or less familiar notions.

Let V indicate the (finite) set of propositional variables $p_1, p_2, \ldots, p_n$ and let us assume the usual formation rules for statements. A constituent is a statement of the form $(\pm)p_1 \;\&\; (\pm)p_2 \ldots \&\; (\pm)p_n$, i.e. each variable occurs negated or unnegated. Hence, an arbitrary constituent C can also be represented by the subset C^+ of those members of V that occur unnegated in C. Conversely, each subset of V corresponds to a unique constituent. By consequence, the powerset $P(V)$ of V represents all possible sets of constituents. From the well-known fact that an arbitrary statement can be written as a disjunction of constituents it now follows that each statement

A can be represented as a subset of $P(V)$:$A^x = \{C^+/C$ is 'constituent-disjunct' of $A\}$. Note that for a constituent C:$C^x = \{C^+\}$.

The members of $P(V)$ are usually called (logically) possible worlds and the standard truth-definition is such that a statement A is true (false) in $w \in P(V)$ if and only if $w \in A^x$ $(w \notin A^x)$.

Using these ingredients, the circuit-example suggests that we need to distinguish a *descriptive* and a *theoretical* point of view in the following way:

D1 The descriptive point of view is directed to one (*the*) *actual world* w_a.

T1 The theoretical point of view is directed to *the set of physically possible worlds*, or *the physical space*, W_{ph}.

D2 A statement A is called *descriptively true* (D-true) if A is true in the actual world, i.e. $w_a \in A^x$, and D-false otherwise.

T2 A statement A is called *theoretically true* (T-true) if A is true in all physically possible worlds, i.e. $W_{ph} \subseteq A^x$, and T-false otherwise.

D3 A statement A is *strong-descriptively true* (Ds-true) if A is true in the actual world and false in all others, i.e. $A^x = \{w_a\}$, and Ds-false otherwise.

T3 A statement A is *strong-theoretically true* (Ts-true) if A is true in all physically possible worlds and false in all physically impossible ones, i.e. $A^x = W_{ph}$, and Ts-false otherwise.

At this point a digression on modal terminology is required. Although our approach will be based on a non-modal (propositional) language it should be stressed that this is only possible by introducing new kinds of truth-values, in particular: T-true/T-false. The present exposition can be translated in modal terms in at least two ways, a standard and a non-standard one. In the standard way we add modal operators to the language (e.g. $\Box A$, to be read 'physically necessary A') and equate: $\Box A$ is true if and only if A is T-true. In this way we arrive essentially at what is called an S_5-system. In the non-standard way we do not 'modalize' the language but the truth-values and equate: A is *physically necessary true* if and only if A is T-true. In the present article we have chosen however for the technical term 'T-true' for three reasons. *First*, it will become clear that there is a coherent intuitive way of speaking about true and false laws and theories, i.e. without using modal terms explicitly (except, of course, the notion of physically possible world). E.g. according to T7 we get: A is a true law iff it is a T-true (T-)statement. *Second*, there does not seem to be a standard modal operator which corresponds to the stronger truth-value Ts-true, defined in T3, which we will need so much. *Third*, a full-fledged plea for the suggested non-standard modal terminology (i.e. modal truth-values) could be postponed in this way to another occasion.

The following statements are immediate consequences:

D4 Apart from equivalent formulations, there is just one Ds-true statement, being a constituent, indicated from now on by D, which may be called *the true description*, or *the descriptive truth*. Note that $D^+ = w_a$ and that $D^x = \{w_a\}$.

T4 Apart from equivalent formulations, there is just one Ts-true statement indicated from now on by T, which may be called *the true theory*, or *the theoretical truth*. Note that $T^x = W_{ph}$.

Hence, there are essentially two problems of verisimilitude:

PDV *The problem of descriptive verisimilitude*: explicating the idea that one statement may be closer to the descriptive truth than another.

PTV *The problem of theoretical verisimilitude*: explicating the idea that one statement may be closer to the theoretical truth than another.

Of course, separating the two problems does not imply that their solutions need to be different. After all, both problems concern the comparison of three statements. However, different solutions might be expected if the statements to be compared in the one problem would be of a different nature as those to be compared in the other problem. There is indeed a distinction between statements, that obviously can be important in this connection. It is the following distinction between descriptive and theoretical statements:

D5 A statement is a *descriptive statement* (D-statement) if it is a conjunction of different negated and unnegated propositional variables (hence, a constituent is a (complete) *D*-statement).

T5 A statement is a *theoretical statement* (*T*-statement) if it does not imply any *D*-statement (which implies that a *T*-statement is not complete).

Note that this distinction is not exhaustive: there is room for still another type of statements, to be called *mixed* statements: although p_1 & $-p_2$ is a *D*-statement and $p_3 \rightarrow p_4$ is a *T*-statement, their conjunction is of neither type. Note also that, in the circuit-example, D_c is a (complete) *D*-statement, i.e. a constituent, whereas T_c is a *T*-statement.

The way in which the descriptive point of view has been formulated implies that the descriptive truth is a constituent (see D4). However, the theoretical point of view as formulated up till now does not imply that the theoretical truth is not a constituent, let alone that it is a *T*-statement. This means that there is room for both a weak and a strong additional assumption within the theoretical viewpoint:

WTA *The weak theoretician's assumption*: there is more than one physically possible world, i.e. the theoretical truth is not a constituent.

STA *The strong theoretician's assumption*: the theoretical truth is a *T*-statement.

Note first that in a context where WTA is not satisfied, i.e. where there is just one physically possible world (the actual one) all our distinctions collapse. Hence, in this case there is no separate room for the theoretical point of view differing from the descriptive point of view. We will call such a context *a non-theoretical context*.

On the other hand, a context satisfying STA will be called *a theoretical context*. Finally, the intermediate case of a context satisfying WTA but not STA might be called *a semi-theoretical context*. Without loss of generality we may leave semi-theoretical contexts out of our considerations. For, in such a context the theoretical truth is either an incomplete D-statement (and hence not 'containing' all variables) or a mixed statement. In the first case, the context becomes essentially a non-theoretical one, if we restrict our language to the relevant variables. In the second case, where the theoretical truth is a mixed statement, it will be possible to rewrite it as a conjunction of an incomplete *D*-statement and a *T*-statement with non-overlapping propositional variables. In the light of the first case we may conclude that this reformulation leads to a separation of the original context into two subcontexts, viz. a non-theoretical one and a (genuine) theoretical one.

Of course, the fact that in a theoretical context the theoretical and the descriptive point of view are essentially distinct does not imply that they are not relevant to each other. On the contrary, both points of view are of mutual interest just because the actual world is one of the physically possible worlds.

Hence, from now on we will presuppose a theoretical context and investigate the two points of view and their interaction.

Returning to the two problems of verisimilitude (PDV and PTV) it is obvious that the distinction between *D*- and *T*-statements (and mixed ones) may have consequences for their solutions, in the sense that these solutions might be restricted to comparisons of *D*-statements and *T*-statements, respectively.

In the next section it will be shown that there is a general (unrestricted) solution to PTV, i.e. including the comparison of *D*-statements. In Section 4 we will see that the solution of PDV is restricted to *D*-statements in a natural way. In Section 5 we will relate these solutions to the discussion about verisimilitude in the literature.

We conclude the present section with plausible explications of the informal terminology used at the beginning of this section and some other terminology as well, using the introduced distinction between *D*- and *T*-statements and the introduced truth-notions.

D6 A statement is called a factual, descriptive or basic statement if and only if it is a *D*-statement.

T6 A statement is called a law, lawlike statement or theory if and only if it is a *T*-statement.

Although T6 may well be defended, we must confess that it is in many theoretical discussions more common practice and sometimes also more convenient to call all statements laws and theories. If we do so we will speak of *the unrestricted terminology* of laws and theories.

However this may be, in the following explications of truth-terminology we will see that the truth-status of a statement may vary with the way it is called.

D7 A statement (law or theory) is called true/false (actually true/false) if and only if it is *D*-true/*D*-false.
A constituent is called the true/a false description if and only if it is *Ds*-true/*Ds*-false.

T7 A law is called true/false if and only if it is *T*-true/*T*-false.
A theory is called the true/a false theory if and only if it is *Ts*-true/*Ts*-false.

According to these explications a (*T*-)statement may be a true statement, but a false law and hence a false theory, but it may also be a true statement, a true law, but a false theory. These modes of speech are obviously related to expressions like 'actually true, but false as law', as may be said, under proper circumstances, of e.g. 'if it rains then it is night'.

It might be objected to the explication T6 that it implies, in view of T5, that a simple disjunction like $p \vee q$ is already called a law. The problem here is that the law-terminology is ambiguous in the following sense. In practice we talk about true and false laws as well as about laws simpliciter. In the latter case, say '*A* is a law', we frequently imply that *A* is true (in some sense, here explicated as *T*-true). But the former mode of speech, in particular '*A* is a false law', implies that it must also make sense to say '*A* is a law', without implying that it is true. (Note that '*A* is a theory' is always used in this sense.) Of course, the explication in T6 of '*A* is a law' should be understood in this weak sense, i.e. as equivalent to '*A* is a lawlike statement'.

The suggested alternative explication, let us call it 'a law in the strong sense', is of course: a statement is a law in the strong sense iff it is a *T*-true *T*-statement, otherwise it is not a law in the strong sense.

By consequence, although $p \vee q$ is a law (in the weak sense) according to T6, it is only a *true* law, according to T7, if it is in addition *T*-true. In this case, and only in this case, it is a law in the strong sense.

But one might still object that calling a statement a law does not only imply that it is true in our, essentially context-dependent, sense but that it is true in some universal, context-independent sense, as is even more the case when we call a statement a *law of nature*. To be sure, we do not pretend to have explicated this universality aspect which might be associated with the law-terminology.

On the contrary, a main conclusion of the exposition up till now is that the terminology of laws (and theories etc.) can be relativised in a coherent way to a certain (theoretical) context. In the following sections we hope to show that it is also a fruitful step. But the problem of explicating laws of nature, as distinguished from context-dependent laws, is left open; only the discussion at the end of Section 5 will be of some interest to this problem.

3. THEORETICAL VERISIMILITUDE

As we announced earlier, there is a general solution to the problem of theoretical verisimilitude and hence we will use the unrestricted terminology of laws and theories, but of course stick to their respective truth-conditions (T7).

Let us start from the Popperian intuition that a law is false if and only if it has counterexamples, whether already discovered or not. If we combine this with our definition that a law A is false iff it is T-false, i.e. $T^x \nsubseteq A^x$, it follows that the notion of a (physical or *real*) counterexample to a (putatively true) law A refers to a physically possible world in which A is false, i.e. which is not included in A^x. However, it also follows from our definitions that, although all true laws are logical consequences of the true theory, which is the strongest true law, most of them are false theories. This suggests to call a physically *im*possible world in which a (putatively true) theory is true a non-physical, or *virtual*, counterexample to that theory.

The set of all (real and virtual) counterexamples to a theory A is, according to these definitions, the symmetric difference between A^x and T^x, i.e. $A^x \Delta T^x =_{\mathrm{df}} (T^x - A^x) \cup (A^x - T^x)$, see the two shaded areas in Figure 2. If it is non-empty then A is a false theory.

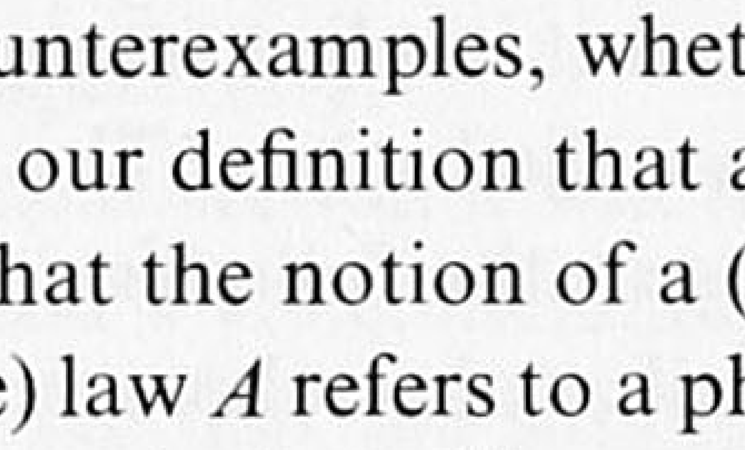

Fig. 2.

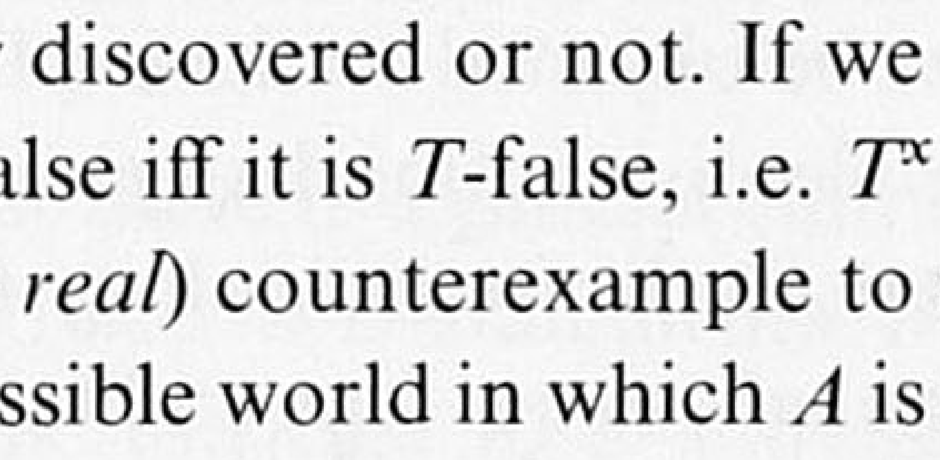

Fig. 3.

Recalling that in the present context all statements are called theories the following definition is now the quite plausible and unrestricted solution to the problem of theoretical verisimilitude.

DEFINITION 1. A theory B is closer to the theoretical truth T than A ($A <_T B$) iff $B^x \Delta T^x \subset A^x \Delta T^x$. ($A \leq_T B$ is defined by replacing '$\subset$' by '$\subseteq$'.)

In Figure 3, $A \leq_T B$ amounts to the emptiness (on logical grounds) of the two double-shaded areas: A has at least all counterexamples that B has. Using informal terminology, it follows from Definition 1 that it is very well possible for a false theory to be closer to the truth than another false theory.

It is also easy to check that $\leq_T$ and $<_T$ are (reflexive resp. irreflexive) partial orderings on the set of all statements, for fixed T of course. Hence, the solution leaves room for the idea of a sequence of false theories converging to the truth.

So we see that Definition 1 satisfies two central adequacy conditions: the applicability to false theories, and the possibility of convergence to the truth.

Also if applied to laws Definition 1 has some attractive consequences. A true law may be closer to the truth than another true law. The same holds for false laws. Moreover, a true law may be closer to the truth than a false law, but a false law cannot be closer to the truth than a true law.

Finally, it is easy to check that in the case of two true laws (but false theories), neither of which implies the other, the conjunction is closer to the truth than either of the conjuncts. This is in agreement with the intuition that we know more about the truth if we find new true laws.

4. DESCRIPTIVE VERISIMILITUDE

In dealing with the problem of descriptive verisimilitude we shall first restrict our attention to constituents, the descriptive truth D being one of them. The following step then consists in the generalization to the case of descriptive statements.

Hence, our first problem is to explicate the idea that one constituent B may be closer to D than another constituent A. Let us return for a moment to the circuit-example, and let us compare the following constituents:

(A_c). p_1 & p_2 & $-p_3$ & p_4 & q;
(B_c) p_1 & $-p_2$ & $-p_3$ & p_4 & q;
(D_c) p_1 & $-p_2$ & p_3 & p_4 & q.

Whereas D_c gives the true description, both A_c and B_c give false descriptions. Nevertheless we are inclined to say that B_c is closer to D_c than A_c, in the plausible sense that B_c makes less mistakes.

This idea is easily generalized, even in a way similar to our treatment of theoretical verisimilitude. A mistake made by a constituent A is simply a propositional variable about which A and D disagree. Recall that A can be represented as the subset A^+ of V of the unnegated variables occurring in A. Hence, the set of mistakes made by A is the symmetric difference between A^+ and D^+, i.e. $A^+ \Delta D^+ =_{\text{df}} (D^+ \text{-} A^+) \cup (A^+ \text{-} D^+)$. This leads us to the following definition as the plausible solution to the problem of descriptive verisimilitude restricted to constituents.

DEFINITION 2^C. A constituent B is closer to the descriptive truth D than constituent $A(A <_D B)$ iff $B^+ \Delta D^+ \subset A^+ \Delta D^+$. ($A \leq_D B$ is defined by replacing '$\subset$' by '$\subseteq$'.)

Of course, we could draw similar pictures as Figures 2 and 3, with adapted interpretations. Here also it may very well occur that one false description is closer to the truth than another. Moreover, $<_D$ and $\leq_D$ are partial orderings on the set of constituents, and hence on $P(V)$. By consequence, a sequence of false descriptions converging to the truth is very well possible; an actual example, to which this definition is applicable is e.g. an historical account in which successive corrections are made.

For the generalization of Definition 2^C from constituents to descriptive (D-)statements we start again with a circuit-example. Compare the descriptive statements

(A'_c) $-p_1$ & p_2 & p_3
(B'_c) $-p_2$ & $-p_3$ & p_4 & q
(D_c) p_1 & $-p_2$ & p_3 & p_4 & q

Intuitively, we want to arrive upon a definition according to which B'_c is closer to D_c than A'_c because, as far as they 'overlap', B'_c makes less mistakes, while in the region where they do not overlap, A'_c makes only mistakes, whereas B'_c makes no mistakes. Note that any weaker condition for the non-overlapping parts would require the assignments of weights to mistakes.

According to these intuitions, the plausible solution to the problem of descriptive verisimilitude restricted to descriptive statements is (using $V(A)$ to indicate the variables occurring, negated or unnegated, in the D-statement A):

DEFINITION 2. A descriptive statement B is closer to the descriptive truth D than the descriptive statement $A (A <_D B)$ iff

(i) $(B^+ \Delta D^+) \cap (V(A) \cap V(B)) \subseteq (A^+ \Delta D^+) \cap (V(A) \cap V(B))$;
(ii) $(B^+ \Delta D^+) \cap (V(B)\text{-}V(A)) = \emptyset$;
(iii) $(A^+ \Delta D^+) \cap (V(A)\text{-}V(B)) = V(A)\text{-}V(B)$;

with the additional clause that '$\subseteq$' in (i) should be replaced by '$\subset$' if $V(A) = V(B)$. ($A \leq_D B$ is defined by deleting the additional clause).

The additional clause may need some explanation. If we had required in (i) proper inclusion in general the definition would not be prepared for cases where A and B make the same mistakes as far as they overlap but where there are non-overlapping variables satisfying the above mentioned intuition. And we want, of course, to have, e.g., that p & q is closer to p & q & r than p and even that $-p$ is closer to p & q & r than $-p$ & $-q$.

Now note first that Definition 2 reduces to Definition 2^C if A and B are constituents (if we substitute $V(A) = V(B) = V$, (ii) and (iii) become trivially satisfied). Restricting our attention to D-statements, it follows from Definition 2 that a false statement (i.e. a D-false statement) may, informally speaking, be closer to the truth than another false statement, and the same holds for two true statements. Also, a true statement may be closer to the truth than a false one. But, it is not difficult to check that on the other hand a false statement cannot be closer to the truth than a true one.

Furthermore, the conjunction of two true statements, neither of which implies the other, is closer to the truth than either of the conjuncts, in agreement with the intuition that we know more about the truth if we have found new true statements.

Finally, the generalized relations $<_D$ and $\leq_D$ remain partial orderings, now on the whole set of descriptive statements. Again, by consequence, a sequence of false descriptive statements converging to the truth is possible.

Although the solutions to the two problems of verisimilitude are different, they are structurally very similar. It is even possible to reconstrue Definition 1 as an application of Definition 2^C as we shall show now.

The statements '$w \in T^x$' and '$w \notin T^x$', where $w \in P(V)$, may be interpreted as elementary propositions: the world w is physically possible, the world w is physically impossible, respectively. We might call

$$(\mathrm{d_A}) \qquad \&_{w \in A^x} (w \in T^x) \;\&\; \&_{w \notin A^x} (w \notin T^x)$$

the descriptive claim associated with the theory A. Now it is easy to prove that theory B is closer to T than A (in the sense of Definition 1) if and only if $\mathrm{d_B}$ is closer to $\mathrm{d_T}$ than $\mathrm{d_A}$ (in the sense of Definition 2^C).

From this presentation, we see that the idea of theoretical verisimilitude corresponds to that of descriptive verisimilitude on the meta-level. For this reason one might say that the descriptive notion is the more fundamental one.

But there are also structural differences. A first one amounts to this. Comparing Definition 1 and Definition 2 we see that, although both definitions treat the general cases, only Definition 2 has a naturally restricted variant (Definition 2^C) for the special kind of D-statements called constituents. Even the analogue of such a special kind of statements on the level of T-statements does not seem to exist. However, it might also be that Definition 1 should only be compared with Definition 2^C and that it can still be generalized to some kind of higher-order statements, e.g. 'partial' conjunctions suggested by $\mathrm{d_A}$, i.e. conjunctions not exhausting $P(V)$.

A second structural difference is that the solution of the problem of descriptive verisimilitude is restricted to D-statements, whereas the 'theoretical solution' is unrestrictedly applicable to all statements. This restriction to D-statements in the descriptive case is fundamental and natural, as can be easily seen. The 'descriptive solution' is fundamentally based on the notion of a mistake, i.e. a propositional variable about which a descriptive statement and the descriptive truth disagree. But a T-statement cannot make mistakes in this sense, since this would come down to implying at least one D-statement, which is forbidden by our definition of a

T-statement. Hence, given the central idea of the descriptive solution, its restriction to *D*-statements is quite natural.

Conversely, however, recalling that the central notion of the theoretical solution is that of a counterexample, it is easy to see that a *D*-statement may have, formally speaking, a counterexample, notwithstanding the fact that it may sound odd in particular cases. E.g., in the circuit example the actual world is a counterexample to p_1 & p_2. But despite the unnaturalness of this mode of speech there is no formal restriction to apply the theoretical solution to *D*-statements.

5. COMPARISONS

With a number of qualifications one may say that Definition 1 was given by Miller (1978) and Definition 2^C by Niiniluoto (1978). The most important qualification to these attributions is that Miller and Niiniluoto, as well as Tichý (1978) in his approach, identify the truth with descriptive truth.

Hence, to obtain Miller's definition from Definition 1 we have to replace *T* by *D* in it.

Unlike Miller, Niiniluoto needs to generalize for his purpose Definition 2^C to all statements (and not only to *D*-statements), but there is no need to present that generalization here.

Tichý's definition is, unfortunately, very complicated, but it seems that, if restricted to (propositional) constituents, his definition coincides with Definition 2^C.

Unfamiliar with Miller's article we reached Definition 1 by starting the investigation of verisimilitude from the conviction (in the line of the structuralist approach) that empirical theories are not, and should not be, complete and hence that theoretical scientists are not, or at least not only, aiming at the descriptive truth.

Apart from the identification of the truth with the descriptive truth the three approaches mentioned are in two technical respects more general than ours. In the first place, they presuppose a first-order language instead of a propositional language. The following remarks are relevant here. Although it is technically somewhat involved there are no essential difficulties in generalizing Definition 2^C to first-order complete theories, as is clear from Niiniluoto (1978). Moreover, it is quite easy to extend Definition 1 to first-order theories: $P(V)$ can simply be replaced by the class of structures of the relevant first-order language.

In the second place all three approaches mentioned include non-qualitative or non-deductive considerations in one way or another, assuming weights for counterexamples and mistakes. With respect to this we confine ourselves here to the remark that weights will always have some arbitrary aspect and hence it is not of primary importance to introduce weights.

Despite the above qualifications to the attributions, we may conclude that the two main solutions (Definition 1 and Definition 2^{C}) were already present in the literature. But we have also to conclude that there is a fundamental confusion in the literature, stemming from the conflation of two different problems of verisimilitude. And, on its turn, this conflation of problems has arisen from the neglect of the theoretical point of view, as we have called it.

We hope that in presenting the *two* problems and their solutions, we have made a convincing case for the thesis that the theoretical point of view should be taken into consideration in dealing with 'the' problem of verisimilitude. But we think to have shown more reasons for the introduction of this theoretical viewpoint. For at the same time it rendered explications of much informal terminology currently in use, especially concerning laws, theories and counterexamples.

In Appendix 1 we will compare Definition 1 with Popper's original definition and see what was wrong with it apart from the identification of the truth with the descriptive truth. In Appendix 2 we will compare the theoretical solution with the descriptive one for the case where they are formally competitive, viz. in a non-theoretical context where $T = D$. It turns out that the solutions agree in this case always about true statements. In Appendix 3 we will compare the present account with the plea for 'legisimilitude' by Cohen (1980) which is not discussed in Oddie's paper. Although quite similar in its critical point, it will be argued that Cohen's constructive claim is fundamentally weaker than ours.

As we remarked already, Miller's definition can be obtained from Definition 1 by replacing T by D in it. Formally we can reproduce this in our framework by assuming a non-theoretical context, where the theoretical truth boils down to the descriptive truth. The relation 'closer to T' ($<_T$) according to Definition 1 will be indicated by $<_{T=D}$ for such a context. Now it is interesting to discuss a standard counterargument, which happens to be applicable to $<_{T=D}$ and hence to Miller's definition. E.g. Oddie (1981) devotes a lot of attention to it.

It is easy to check that $<_{T=D}$ has the property that, of two (D- = T-) false statements A and B, B is closer to the truth than A if and only if B is stronger than A (i.e. $B^x \subset A^x$). Now it is argued that this property makes it child's play to approach the truth: take any false statement A and an arbitrary statement B (not implied by A), then the statement A & B is closer to the truth than A.

Although we do not dispute the argument as such, we completely disagree about its evaluation. In a non-theoretical context, where we are looking for the descriptive truth, which is a constituent, the end-product of our descriptive activities will also be a constituent. The described child's play, starting from a false statement, will end in one or other constituent different from D, and hence it will not only be just a false statement but a false description. It is easy to check that of two different false descriptions the one cannot be closer to the truth T (= D) than the other in the sense of Definition 1. Hence, the evaluation of the indicated, descriptive truth-play depends entirely on whether or not there is a solution to the first, restricted, problem of descriptive verisimilitude we have met: (in suggestive terms) the explication of the idea that one false description may nevertheless be closer to the true description in some other sense than Definition 1. Of course, in our opinion Definition 2^C provides the solution to this problem, and hence the descriptive truth-play can be played better or worse.

For theoretical contexts it is even more easy to show that the above argument is without danger. The analogue of the problematic situation is now of course that A and B are false laws (T-false) and that B is stronger than A. But now Definition 1 leads only generally to the conclusion that B is closer to T than A if A, and hence B, is incompatible with T. In other words, only if A has only physical counterexamples or, one might say, only if A does not contain a grain of truth, the child's play can be played. But in this case even the contradiction is closer to the truth than A, in other words, following the path from A to B soon learns that we have to start from scratch. Of course, in this argument we do not take into consideration the fact that in real science a theory may well contain not a grain of truth in the indicated logical sense and still be very close to the truth in some sense of approximation. But this kind of approximation clearly falls outside the scope of the present article.

We conclude this section by considering the possibility that there might still be good reasons to neglect the theoretical point of view in the discussion about verisimilitude.

In the first place one might argue that the concept of physically possible world is rather vague. This may be admitted, but is the concept of actual world really less vague? Moreover, the explicative results, in the previous sections and the following ones, show that the concept of physically possible world is essentially involved in much informal terminology concerning science.

A second, more specific, argument in favour of the neglect of the theoretical point of view is more interesting. In the literature the problem of verisimilitude is usually presented as the problem of explicating and justifying the idea that science can approach the true description of the actual world in the literal sense of our whole universe, including its past and its future. This might be called the problem of universal descriptive verisimilitude. The natural question now is whether there is also a problem of universal theoretical verisimilitude, i.e. does it make sense to talk not only about the actual universe and the universal descriptive truth but also about physically possible universes and the universal theoretical truth? Or, to put it differently, is the universal context a theoretical one, or a non-theoretical one?

At first sight it might seem strange to hesitate in answering that the universal context is theoretical. For, one may argue, that this is so is clear from our exposition that there are many restricted contexts which are evidently theoretical, e.g. the circuit-example. If we combine such restricted contexts we will get the descriptive truth about the combined context by forming the conjunction of the component descriptive truths. Now we might expect to arrive at the theoretical truth about the combined context in the same way. But this is not true. Although the conjunction of theoretical truths will lead to a true law about the combined context, it need not be its true theory. For it may well be that two physically possible worlds belonging to different restricted contexts cannot coexist.

E.g. in the circuit-example the represented actual world, in which the bulb lights, is of course physically possible. In our treatment of the circuit we did not include the two physically possible worlds where the source of voltage is in or out. Now, in the combined context any combined world in which the source is out and the bulb lights is of course physically impossible, despite the fact that it is trivially allowed by the conjunction of the restricted theoretical truths.

Hence, in general, enlarging the context may lead to stronger theoretical truths than the conjunction, but it can of course not lead to weaker theoretical truths.

The crucial question now is where the process of combining contexts up to the universal context brings us if we start from restricted contexts and the corresponding theoretical truths. Does it bring us in the end to the universal descriptive truth, i.e. to the same endpoint as the one we would reach if we had started from the corresponding descriptive truths? If so, the universal context is a non-theoretical one, otherwise it is theoretical.

We do not have a clear opinion about the answer and we will confine ourselves to three remarks. First, even if the ideal of one unified fundamental theory in physics would have been realized, then it would still not be a complete theory. That is, that unified theory would, like all other theories, need additional initial conditions to explain or predict individual events. Second, if nevertheless the universal theoretical truth coincides with the universal descriptive truth there must have been during the described combination-process transitions from theoretical statements to descriptive statements, and it is not easy to understand how this is possible. Third, if the universal context is non-theoretical there is, after all, just one problem of *universal* verisimilitude, the descriptive one and Definition 2 provides its solution.

However this may be, if we are interested in an adequate picture of science we have not only to take seriously that the aims of scientists are more modest than the universal truth but also that their aims are therefore essentially two-fold: descriptive truths and theoretical truths, leading to two essentially different problems of verisimilitude as we have seen.

6. EXPLANATION AND PREDICTION

The distinction between the descriptive and the theoretical point of view, especially the respective truth-notions and the distinction between descriptive and theoretical statements, sheds also new light on the discussion about explanation and prediction. In the literature (e.g. Nagel (1961)) there is general agreement that the explanation and prediction of individual events should be based, if possible, on a so-called deductive-nomological argument. This argument-form presupposes some distinction between lawlike (or theoretical) statements, expressing the required law- or theory-premiss(es), and descriptive statements, expressing the so-called initial conditions and the event to be explained or predicted. Both kinds of premisses are usually supposed to be needed for the deduction.

In the literature there is also more or less general agreement that lawlike statements should be construed as universally quantified conditional statements. This, however, excludes the possibility of lawlike statements within a propositional language. Note, by the way, that the attempts to find a characterization of lawlike statements is in contrast with the easy way of talking about laws and theories in the literature about verisimilitude.

From our exposition of a theoretical context it is almost evident what the (formal) requirements for an adequate deductive-nomological argument are, if we restrict our attention to a propositional theoretical context:

– there is a, perhaps complex, T-statement L and there are, perhaps complex, D-statements C and E, such that:
– L is T-true and C is D-true,
– 'if L and C then E' is a logically valid argument,

If these requirements are fulfilled E is of course D-true.

What about the usual requirement that both types of premisses should be needed in the derivation of the conclusion? Of course, E being a D-statement and L a T-statement it is impossible that L implies E, by definition. On the other hand, C implies D-statements on its own. Hence, the necessity to use a T-statement as additional premiss follows only if we require also

– 'if C then E' is logically invalid.

This additional requirement may also be interpreted as the explication of the idea that the argument may not be circular.

The argument thus construed reflects typically the interaction between the descriptive and the theoretical point of view. In terms of worlds, using some information about the actual world (C) and about the class of physically possible worlds (L) it is possible to derive new information about the actual world (E).

In the circuit-example an argument satisfying all requirements is, e.g., L: $(p_3 \,\&\, p_4) \rightarrow q$, C: $p_3 \,\&\, p_4$, E: q.

Of course, if we have an argument satisfying all requirements, except perhaps the truth-conditions, and if E happens to be (D-)false there are two non-exclusive possibilities: L may be D-false and hence it is T-false, or C may be D-false. This is in perfect agreement with the general intuition that explanation and prediction may fail for two quite different reasons: descriptive and theoretical ones.

It is interesting to note that the descriptive truth can never be used as a premiss in an adequate deductive-nomological argument, for it is a complete D-statement and hence it can only give rise to circular arguments. Of course, we take this aspect of the descriptive truth as (additional) evidence for the thesis that scientists are not only aiming at the descriptive truth, especially when they want to explain and predict.

As to the explanation and 'prediction' of laws we get as adequacy-conditions for the argument:

– there are, perhaps complex, T-statements L and L', such that
– L is T-true,
– 'if L then L'' is a logically valid argument,

If these conditions are satisfied L' is of course T-true.

Here again we are inclined, in order to exclude circular arguments, to require in addition:

– 'if L' then L' is invalid;

i.e., we want to exclude 'self-prediction/-explanation'. It is also frequently argued that the underlying argument may not be too easy in the sense that L is simply stronger than L', but should consist of at least two independent laws. However, even if we take independency simply in the sense that neither law implies the other, there is no natural place for this requirement in our approach, which suggests that it is an intelligible but psychological requirement.

In the circuit-example an argument satisfying all requirements is, e.g., L: $q \to (p_4 \;\&\; (p_1 \vee p_3))$, L': $q \to p_4$.

Of course, if we have an argument satisfying all requirements, except perhaps the truth-condition, and if L' happens to be T-false (suppose a counterexample has been found) we may conclude that L is also T-false. But if L is complex it may of course be the case that only one component-law is T-false. If the argument is adequate in all respects the next problem is to explain L. It is clear that we can proceed this kind of explanation of laws up to the point where we have reached the strongest true law, i.e. the true theory. Here we will not go into the problem of explaining the true theory (in some theoretical context), but only remark that the present account suggests that there is a natural transition from the explanation of true laws to that of *the* true theory.

Note that the true theory, being the strongest true law, explains, according to our requirements, all true laws, except itself. This is in perfect agreement with what we intuitively expect. Moreover, we not only have that all logical consequences of the true theory are T-true, but also that they are T-statements, which is a direct consequence of the fact that the true theory is supposed to be a T-statement. Hence, all logical consequences of the true theory are true laws in the strict sense of T-true T-statements.

We conclude from the foregoing exposition that the introduced truth-notions and the distinction of D- and T-statements provide the adequate means for explicating intuitions about explanation and prediction. Moreover, we have the feeling that it will be worthwhile to reconsider a number of other traditional problems in their light. We think of the following problems in particular. (1) The realism-instrumentalism-debate, as far as the notion of truth is concerned. (2) The debate about the correspondence and the coherence theory of truth. (3) The debate about the possibility of 'individual' laws and theories, i.e. laws and theories about one particular system, e.g. the circuit. (4) The analysis of dispositional terms and of subjunctive and counterfactual conditionals. (5) The analysis of physical modalities of statements, such as physically necessary/possible etc. (6) The analysis of methodological rules.

Leaving the first five problems for other opportunities we will restrict our attention in the rest of this article to the sixth problem.

7. METHODOLOGICAL RULES

7.1 The Rule of Success and the Rule of Content

In the foregoing sections we assumed to know, when necessary, the true theory T. Doing (theoretical) science presupposes, however, not to know T but to aim at T. In this section we will investigate some methodological rules that may govern the search for the true theory in a theoretical context. Although the exposition will be formally restricted to propositional theories, much of it can easily be extended to first-order theories.

Again we can use the unrestricted terminology of theories: all statements may be called theories. But, recall also, that if a statement is called a theory its corresponding truth-condition is Ts-true, which implied that almost all theories are false.

In Popperian spirit we call a theory A *empirical* if it forbids something, i.e. if $P(V)$-A^x, called the *empirical content* of A, is not empty. A^x itself may be called the *proper content* of A. In Section 3 we called already the members of T^x-A^x the real counterexamples of A and those of A^x-T^x its virtual counterexamples. By analogy we may call the members of $T^x \cap A^x$ real examples of A, and those of $\overline{T^x} \cap \overline{A^x}$ virtual examples.

Of course, testing theories is only possible on the basis of physically possible worlds. Hence, in the process of searching the true theory we will have to deal with virtual counterexamples in other ways. With this in mind, an empirical test of a theory A comes down to the following: the experimentor (or nature) realizes some physically possible world $w \in T$, i.e., the actual world of the test, he gives a (complete) description D_w and checks whether D_w and A are in conflict or not.

Neglecting the possibility of descriptive mistakes, and hence of false descriptions, it is of course natural to call the test a *falsification* of A (and w a (real) counterexample) if A and D_w are incompatible; and a *confirmation* (and w a (real) example) if they are compatible. From a falsification we may conclude that A is T-false and hence Ts-false. From a confirmation we may not conclude that A is T-true, let alone Ts-true. Hence, there is an essential asymmetry between falsification and confirmation. Of course, in a finite propositional context it is possible to exhaust the set of physically possible worlds by a finite number of tests, which makes verification even possible. The essential point however is that the asymmetry holds for a single test.

Let E indicate the disjunction of the (true) descriptions of the investigated physically possible worlds E^x at a certain moment. We may call E the experimental results or evidence. The following 'E-relative' notions for theories A and B are obvious.

A and B are *equally successful* $(A =_E B)$: $A^x \cap E^x = B^x \cap E^x$

B is *at least as successful as A* $(A \leq_E B)$: $A^x \cap E^x \subseteq B^x \cap E^x$

B is *more successful than A* $(A <_E B)$: $A^x \cap E^x \subset B^x \cap E^x$

In Fig. 4 we have drawn the situation that B is more successful than A. Note that even B may well be falsified.

Comparing Figure 4 with Figure 3 it is easy to read off that, according to Definition 1:

– it is impossible that A is closer to T than B, and that

– it is still possible that B is closer to T than A.

This suggests the following methodological rule:

(RS) *Rule of Success*: if $A <_E B$ then prefer B to A.

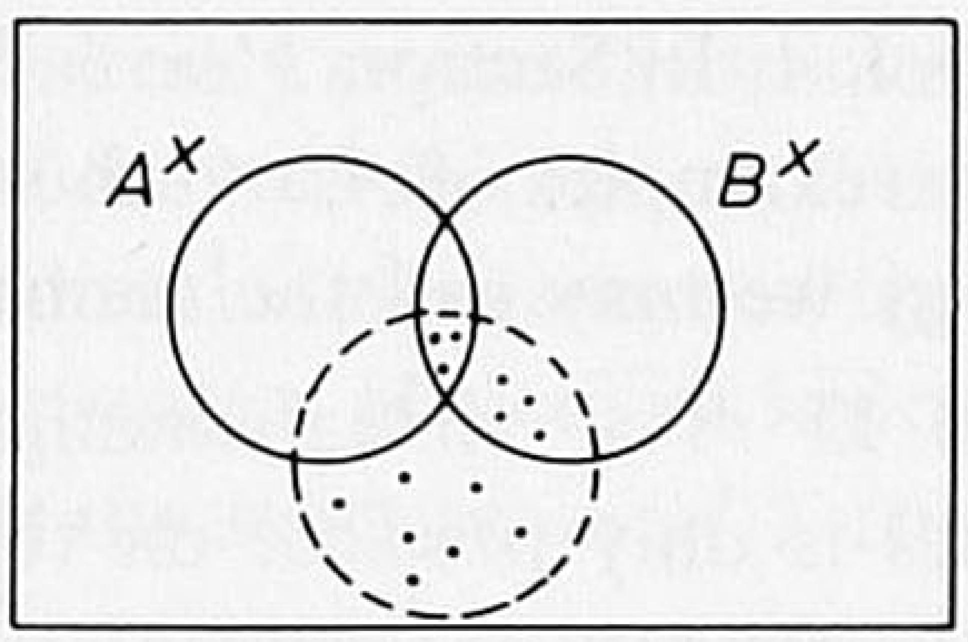

Fig. 4. Points represent members of E^x. The interrupted circle indicates the unknown T.

The indicated justification of this rule does of course not exclude the possibility that new (additional) evidence may destroy the judgement that B is more successful than A, and hence the possibility that B is indeed closer to T than A. But new evidence cannot make A more successful than B (in the defined sense) and hence we cannot arrive at the situation that we have to apply RS in the opposite direction. In other words, we will never have to regret preferences according to RS.

It is interesting to analyse in some detail the situation in which RS is applicable. $A <_E B$ is easily seen to be equivalent to the conjunction of

(a) $B^x - A^x \neq \emptyset$; (b) $A^x \cap E^x \subseteq B^x \cap E^x$; (c) $(B^x - A^x) \cap E^x \neq \emptyset$.

As is well-known Lakatos (1970, p. 116) has derived from Popper's work a notion of 'sophisticated falsification', which reads, with the relevant substitutions:

> ... a scientific theory A is *falsified* if and only if another theory B has been proposed with the following characteristics: (a′) B has excess empirical content over A: that is, it predicts *novel* facts, that is, facts improbable in the light of, or even forbidden, by A; (b′) B explains the previous success of A, that is, all the unrefuted content of A is included (within the limits of observational error) in the content of B; and (c′) some of the excess content of B is corroborated.

With some technical and one principal qualification it is easily seen that the three requirements correspond exactly to (a), (b) and (c). First we strengthen the second part of (a′) to 'that is, it predicts *novel* facts, that is facts forbidden by A'. This clearly excludes the excess-*empirical*-content-interpretation mentioned by Lakatos, i.e. 'B has excess empirical content over A' in the Popperian sense. More precisely, '(empirical) content' should be replaced by '(proper) content' at all places. With these qualifications the correspondence is perfect, except for Lakatos' requirement, contained in (a′) that (some of) the facts corresponding to the excess (proper) content of B should be unknown before the invention of B. In our terminology, this means that it should be unknown that some logically possible worlds in $B^x - A^x$ are physically possible. Although the 'novelty-requirement' is intended to exclude ad hoc theories it is doubtful whether it can serve this purpose for it is obviously not of service in, and may even slow down, approaching the truth. Perhaps, it is possible to explicate the notion of an ad hoc theory in some different way, not using reference to what is (un-)known.

We conclude from this comparison that Lakatos' definition of falsification, except for the novelty-requirement, is the same as the rule of success (RS) and hence that it has a clear justification in the light of approaching the truth.

An obvious objection to RS is, however, the following. For any theory B weaker than A and for any evidence E it holds trivially that B is at least as successful as A. RS prescribes now to prefer any weaker theory B as soon as we find a real counterexample of A which is an example of B. Hence, RS induces so to speak a pressure in the direction of weak theories. Of course, T may or may not be a strong theory itself. Hence, there is no reason to agree with the general rule of Popper to prefer strong theories as such. Nevertheless, some counterpressure, against RS, in the direction of strong theories seems required.

Note that a weaker theory has always at least as many *virtual* counterexamples as a stronger one. Hence, one may also say that RS is likely to introduce more virtual counterexamples and that we need to compensate this by a rule which can reduce the number of virtual counterexamples, without reducing the success. This suggests the following rule:

(RC) *Rule of Content* (or Strength): if B has more empirical content (i.e. is stronger) than A and if $A =_E B$ then prefer B.

It is easy to see that RC cannot be justified directly on the basis of verisimilitude considerations similar to those for RS. For, under the stated conditions in RC, B may indeed still be closer to T than A, but A may also be closer to T than B. The justification for RC comes from the interplay with RS. Assume the conditions of RC, if B is, as a matter of objective fact, not closer to T than A it is possible to produce new evidence E_1 ($E_1^x \supset E^x$) such that A becomes more successful than B. Hence, as soon as we find such evidence, RS applies in favour of A, of course. Moreover, if this has occurred we can never return to the situation of RC in favour of B. In sum, RS compensates the risks which RC prescribes to take in such a way that the same risk will never be taken for a second time. But new risks, according to RC, should of course be taken.

Partition-theories

The motivation for the rule of content was based on the assumption that choice problems between stronger and weaker theories occur frequently. However, in the rest of this section we will concentrate on the situation that the relevant candidates for the true theory are restricted to theories which are not comparable in this way and hence where virtual counterexamples have to be (or have been) eliminated in another way than by RC.

Consider once again the circuit-example. We will assume immediately that the true theory T is of the form: the bulb lights (q) if and only if T_r, where T_r is a statement not using q. Hence, we will restrict our attention to candidates of this form: A: $q \leftrightarrow A_r$.

The same holds in many other cases. E.g. equilibrium theories in physics or economics. Think, e.g., of a balance: the true theory and hence candidate-theories are of the form: the balance is in equilibrium if and only if.... A quite different example is a grammatical theory, it is also of the form: a sentence is grammatical, as judged by native speakers, if and only if it satisfies such and such conditions.

We will call such theories Partition-theories (P-theories) because they lead to two kinds of real counterexamples. E.g., for a grammatical theory, a grammatical sentence *not satisfying* the theory-conditions is a (real) counterexample, but an ungrammatical sentence *satisfying* the theory-conditions too.

In general, for a P-theory A: $q \leftrightarrow A_r$ we will call a real counterexample $w \in T^x - A^x$ a *positive* counterexample if q is true in w (and hence A_r is false in w) and a *negative* one if q is false in w (and hence A_r is true in w). Similarly, we may divide the real examples of A in positive and negative examples.

All this may suggest that everything becomes very complicated, but, fortunately, it is possible to argue that we can leave out virtual (counter)examples completely if the true theory is indeed a P-theory. For some logically possible world w, let w' be the same as w except that the corresponding truth-value of q changes. Note that $w'' = w$. Now, it is easy to see that if w is a physically possible world then w' is physically impossible, and vice versa.

By consequence, if w is a real counterexample to a P-theory A then w' is a virtual counterexample. Similarly, if w is a real example of A, then w' is a virtual example. In sum, a real (counter)example of a theory provides at the same time a virtual (counter)example. In still other words, the elimination of real counterexamples is at the same time elimination of virtual counterexamples and, also, increase of real examples is at the same time increase of virtual examples.

Hence, we can neglect all considerations of virtual (counter)examples (and omit the adjective 'real' for real (counter)examples), which can technically be done in the following way. Let V_r indicate the set of propositional variables $V - \{q\}$; statements using only members of V_r will be indicated by A_r, B_r (and T_r) as suggested before. From the assumption that T is a P-theory it now follows that all restricted logically possible worlds, i.e. the members of $P(V_r)$, are physically possible. Hence, T_r divides the restricted physically possible worlds into those in which q is true and those in which q is false. An arbitrary P-theory A: $q \leftrightarrow A_r$ does this in its own way. In Figure 5 all possible (counter)examples of A are indicated.

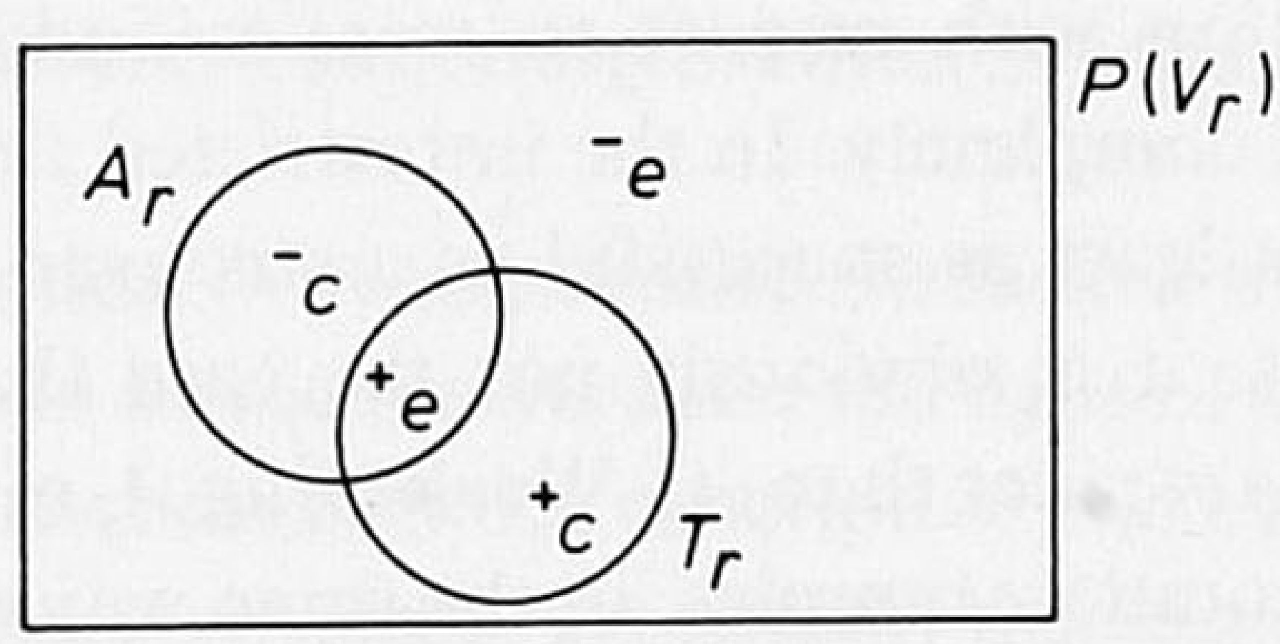

Fig. 5. $+_e$: positive examples of A; $-_e$: negative examples; $+_c$: positive counterexamples; $-_c$: negative counterexamples.

Returning to the evaluation of theories in the light of evidence E, we indicate the descriptions with unnegated q by ^+E and those with negated q by ^-E and their restrictions to V_r by ^+E_r and ^-E_r, respectively. Note that the sets of worlds indicated by them do not overlap. The rule of success (RS) applied when one theory B was more successful than A ($A <_E B$). If A and B (and T) are P-theories this comes about to

(1) $A^x_r \cap {}^+E^x_r \subseteq B^x_r \subseteq {}^+E^x_r$,
(2) $A^x_r \cap {}^-E^x_r \subseteq B^x_r \cap {}^-E^x_r$,

and at least one proper inclusion should hold. In Figure 6 we have drawn the situation that both inclusions are proper.

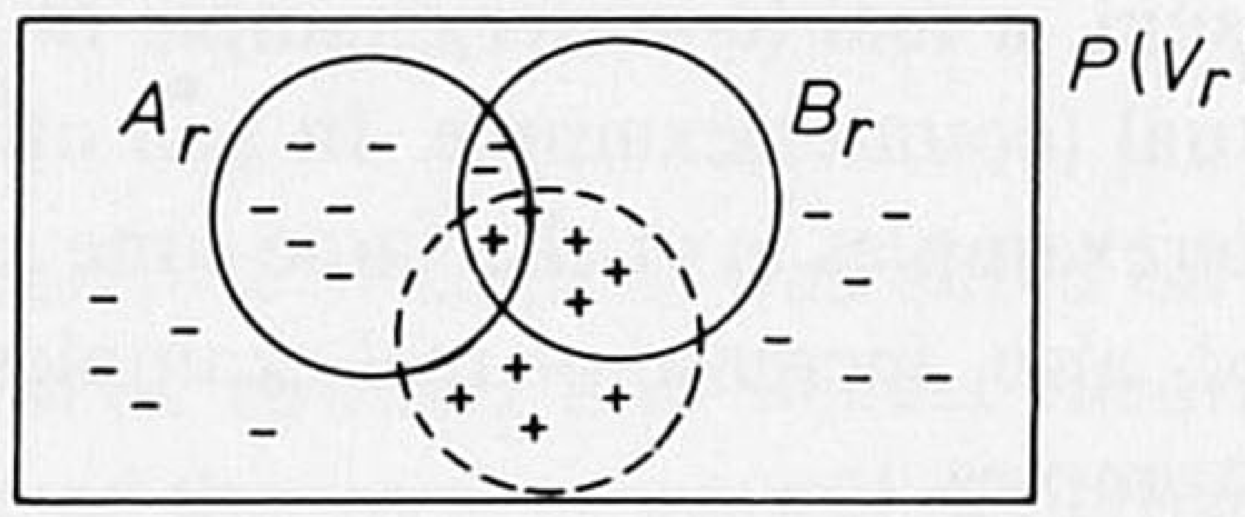

Fig. 6. +-signs represent positive examples of T (${}^+E_r$); −-signs represent negative examples of T (${}^-E_r$). The interrupted circle indicates the unknown T_r.

Note that B is closer to T than A corresponds in the case of P-theories to

(1′) $A^x_r \cap T^x_r \subseteq B^x_r \cap T^x_r$,
(2′) $A^x_r \cap \overline{T^x_r} \subseteq B^x_r \cap \overline{T^x_r}$,

and at least one proper inclusion should hold. The structural analogy with (1) and (2) is again evident.

We may conclude now that in the context of (the restricted versions) of P-theories the situation with respect to weaker and stronger restricted theories has changed completely. In the unrestricted case we saw that any weaker theory was at least as successful as a stronger theory. But, in the context of P-theories, it is obviously not the case that B_r is at least as successful as A_r if B_r is weaker than A_r. Weakening A_r may well increase the number of negative counterexamples, in the same way as strengthening A_r may increase the number of positive counterexamples.

In sum, if we may assume that T is a P-theory, and hence may restrict the attention to P-theories, there is no need, and even no possibility, to compensate the rule of success (RS) by some additional rule.

The reader will be well aware that we have simplified in this section a lot of things, e.g. the exclusion of false descriptions of experimental results and the assumption that we know beforehand what the relevant propositional variables are. In this respect our account of methodological rules is certainly highly idealized. Nevertheless, we hope to have convinced the reader that the presented rules are the proper point of departure for (further) concretization.

Apart from some simplifications we have also restricted our attention to complete (true) descriptions of experimental results. However, it is easy to generalize our account to 'partial descriptions': as soon as there is a logical conflict between a true descriptive statement and the theory, the latter is falsified. This generalization would make it even more clear that descriptive statements, in our technical sense, play the role of Popperian 'basic statements'.

8. FINAL REMARKS

At several places we have already hinted upon the possibility of generalizing our exposition to first-order languages. The most clarifying generalization will perhaps be that of the distinction between descriptive and theoretical statements: it will of course be such that a statement like 'All screws in the car of Mr. Smith are rusty' (to use a favourite example of Nagel (1961)) is a descriptive statement. This explains our intuition that it is not (really) a lawlike statement in a quite different way than the usual diagnosis that it is not unrestrictedly universal.

The generalization to first-order languages will of course introduce a new problem, which is at least of logical interest: the role of non-standard models.

At the end of Section 6 we mentioned already a number of traditional problems that have to be reconsidered in the light of the introduced truth-notions and the distinction between descriptive and theoretical statements. Although most of these problems are usually discussed within a first-order language, we think also here that the main things can already be said within a propositional language.

The same holds in our opinion for the extension of our account to theory-generalization and -specialization. Each theoretical context is supposed to have its own intensional specification of the relevant domain of physical systems. Hence, a larger domain may lead to 'a more general theoretical truth' and a smaller domain to 'a more specific theoretical truth', i.e., the aims of theory-generalization and -specialization.

On the other hand, we think that at least first-order languages are required to extend the analysis to so-called idealization and concretization (i.e., *reculer pour mieux sauter*), where the domain remains essentially constant.

The foregoing remarks may suggest that languages should be taken into account in any case. But it is not difficult to see that, in retrospect, the main things which have been said about theoretical verisimilitude and methodological rules do not need linguistic formulations of theories. That is, the crucial definitions could have been given completely in set-theoretical, hence structuralist, way. From the structuralist point of view the most interesting generalizations of our account are of course the introduction of theoretical terms and constraints. The phenomenon of 'auto-determination' will lead to important qualifications.

To be honest, we did not find examples of real physical theories A, B and T such that B is closer to T than A in the (deductive) sense of Definition 1. We even did not find realistic examples of theories A and B and evidence E such that B is more successful than A (in the light of E). On the other hand, there are of course realistic examples of theories where we would intuitively agree with the two mentioned judgements. Hence, it will be non-qualitative or non-deductive judgements. In order to concretize our account in this respect we will have to introduce distance-measures between worlds allowed by a theory and 'corresponding' physically possible worlds.

Notwithstanding the idealized character of our deductive notions, we hope to have convinced the reader that propositional theories, or even propositional worlds, provide already a model of many informal intuitions and terminology currently expressed and used in the philosophy of science.

Department of Philosophy
State University Groningen

APPENDIX 1: COMPARISON WITH POPPER'S ORIGINAL DEFINITION

In this appendix we will compare Definition 1 with the original definition of Popper (1963, pp. 233–234). We start with writing out $A \leq_T B$ (B is as close to T as A) according to Definition 1 into two parts (i) and (ii) and some equivalent syntactic versions in vertical order. $\vdash$ indicates logical implication, A_L the set of logical consequences of A, and $\bar{X}$ the complement of X.

	(i)	(ii)
(1)	$A^x \cap T^x \subseteq B^x \cap T^x$	$B^x \cap \overline{T^x} \subseteq A^x \cap \overline{T^x}$
(2)	$A \,\&\, T \vdash B$	$B \vdash A \vee T$
(3)	$B_L \cap (-T)_L \subseteq A_L \cap (-T)_L$	$A_L \cap T_L \subseteq B_L \cap T_L$

(iii)

$A^x \subseteq B^x$ or $T^x \subseteq B^x$

$A \vdash B$ or $T \vdash B$

$B_L \cap \overline{T_L} \subseteq A_L \cap \overline{T_L}$

The proofs of the 'vertical equivalences' are not difficult, using the (exceptional) equality of consequence-classes $(A \vee B)_L = A_L \cap B_L$.

Note first that T_L is the set of T-true statements and that $\overline{T_L}$ is the set of T-false statements. Hence, (ii.3) can be read as: all T-true consequences of A are T-true consequences of B. Let us now consider column (iii), that also contains equivalent conditions. We see that (iii.3) can be read as: all T-false consequences of B are T-false consequences of A. It is easy to see that (iii) is stronger than (i) and, hence, that the conjunction of (ii) and (iii) would lead to a stronger definition than Definition 1 gives for $A <_T B$.

However, replacing T by D in (ii) and (iii), or assuming that T is complete, we get precisely Popper's definition, for in this case (ii) can be read informally as: all true consequences of A are true consequences of B, and (iii) as: all false consequences of B are false consequences of A.

For completeness, we will give our version of the knockdown argument of Miller (1974) and Tichý (1974). Suppose (i) and (iii), T complete and B T-false (i.e. $T \nvdash B$). From (iii) and B is T-false we get immediately $A \vdash B$. But, B is T-false and T is complete imply that $T \vdash -B$, hence from (i) we get also $B \vdash A$, and hence A and B are equivalent under the stated conditions. By consequence, (i) and (iii) do not allow non-trivial comparisons of T-false statements, provided T is complete (in which case the T-false statements coincide with the D-false, or simply, with the false statements.)

Of course, without the completeness assumption, (i) and (iii) leave room for non-trivial comparisons of T-false statements, but not as many as (i) and (ii), because, as we remarked already, (iii) is substantially stronger than (i).

Hence, in view of Definition 1, Popper assumed, apart from the completeness assumption, a too strong condition.

We did not find an interesting relation between Popper's definition and Definition 2^C or Definition 2.

One may wonder how it was possible that Popper presupposed the descriptive point of view in the verisimilitude discussion. In particular, as Cohen (1980, p. 500) stresses rightly, because Popper's 'new appendix' on physical necessity (Popper, 1959, pp. 420–441) is a plea for considering physically necessary truths, i.e. in our terminology T-true (T-)statements, as important objectives of science. Cohen adds: 'But he omits to consider its implications for the doctrine of verisimilitude'.

It should, however, also be stressed that Popper has not succeeded in working out the idea of physically necessary truths in a generally accepted way. In our opinion, clarification on this point has long been prevented by the appealing, but highly misleading, suggestion of the terminology of physically possible worlds to the effect that just one of them can be actual (actualized), the other ones being due to *meta*physical speculation.

APPENDIX 2: COMPARISON IN A NON-THEORETICAL CONTEXT

We will compare Definition 1 and Definition 2 for the case of a non-theoretical context, where $T = D$. Of course, the comparison is only possible for D-statements. The relation 'closer to T' according to Definition 1 will be indicated by $<_{T=D}$. Because the distinction between the descriptive and theoretical truth-notions collapses we can simply talk about true and false statements, which means, for D-statements, simply: making no mistakes, making some mistakes, respectively.

It will be convenient to compare the judgements about D-statements A and B according to the four possible cases that A is true/false and B is true/false.

A and B are true. It is easy to check that $A <_{T=D} B$ as well as $A <_D B$ hold if and only if B is stronger than A. Hence, as far as two true statements are concerned, the two judgments always agree.

A is true and B is false. As we already noticed in the respective sections, it can neither be the case that $A <_{T=D} B$, nor that $A <_D B$. But this does of course not yet imply the reverse judgements, which are the object of the next case.

A is false and B is true. Now, $A <_{T=D} B$ holds only in the peculiar case that $B^x - A^x = D^x$. This implies that either B is complete and, hence, that $B = D$, or, if B is incomplete, that any strengthening of B (i.e. with respect to a variable in $V - V(B)$) results in D or a false statement B' implied by A, which implies that $V - V(B)$ contains just one variable about which A must be mistaken. It is not difficult to check that in both cases we have also $A <_D B$. On the other hand, $A <_D B$ holds in many more cases than the peculiar one, but it does not hold in all cases. E.g. it is not the case that p is closer to $p \,\&\, q \,\&\, r$ than $-p \,\&\, q \,\&\, r$.

A and B are false. Here, $A <_{T=D} B$ holds simply if and only if B is stronger than A. On the other hand, $A <_D B$ holds in many cases. If B is stronger than A we have however $A <_D B$ only in the case that B adds only true answers and we have even $B <_D A$ if B adds only mistakes, in all other cases of this kind the relation holds in neither direction.

In sum, we may conclude that $<_{T=D}$ and $<_D$ always agree about true statements, that they never disagree in the strong sense of opposite judgements about a true and a false statement, and, finally, that they may disagree in the strong sense about two false statements.

In the light of the fact that comparing both solutions becomes irrelevant, and even impossible, in a theoretical context the result of our comparison for a non-theoretical context is not unsatisfactory. For, intuitively, there would be a real problem only if the two solutions could disagree in the strong sense about true statements, and this is not the case.

APPENDIX 3: COMPARISON WITH COHEN'S PLEA FOR LEGISIMILITUDE

After finishing this paper I came to understand that Cohen's paper (1980) is in an important sense an anticipation on the present one, despite its fundamentally weaker position.

The critical point of Cohen's essay is directed against, what he calls, the truth doctrine underlying the discussion about verisimilitude. In our terminology this truth doctrine corresponds to the view that science is (only or mainly) aiming at descriptive truth. Cohen gives a number of arguments against this doctrine based on the scientific practice of choosing between hypotheses, of using counterfactual arguments and, hence, of using considerations of physically possible worlds. These arguments provide additional support for our, more artificial, plea against the concentration on descriptive truth and descriptive verisimilitude.

However, the constructive part of Cohen's paper is fundamentally weaker than ours. In fact, Cohen claims that scientists are (also or mainly) aiming at increasing, what he calls, *legisimilitude*, which is a measure for the reliability of a hypothesis as a rule of inference in predictions and explanations.

Restricting our attention to qualitative judgments of legisimilitude we dare to subtract from Cohen's paper, although it is rather informally written, the following definition, using our notation: theory B has at least as much legisimilitude as A if and only if $A^x \cap T^x \subseteq B^x \cap T^x$. It is easy to see, compare e.g. (i.1) in Appendix 1, that this is just half of our definition (Definition 1) of 'as close to the theoretical truth as'. The other half (i.e. ii.1) cannot be added for this would lead to the violation of Cohen's 'consequence-principle' (p. 507) that a consequence of a theory has at least as much legisimilitude as the theory itself.

Hence, from our point of view, Cohen is telling only half of the story. It implies in the first place that he cannot give a logical account of the intuition that scientists can approach the truth; for him there is no *problem* of verisimilitude. It is not for nothing that he calls 'his enemy' the truth doctrine, whereas we would call 'our enemy' the descriptive truth doctrine. In the second place, it implies that Cohen cannot account for certain important activities of scientists, viz. strengthening their theories. In Section 7 we will introduce, and justify, two important methodological rules: the rule of success and the rule of strength. Although Cohen can account for the former, he cannot account for the latter, despite some passages (e.g., p. 504, below) that might suggest the contrary. The crucial point is that there is no reason to strengthen theories for, due to the consequence-principle, this cannot increase legisimilitude.

It is of some interest to evaluate an argument of Cohen against the truth doctrine (pp. 494–499). He argues, quite convincingly, that a measure for the 'scientific merit' of hypotheses should satisfy the following 'conjunction-principle': the merit of the conjunction of two logically independent hypotheses has to be lower than that of the conjunct with the highest merit.

Unfortunately, I do not think that Cohen is right in his claim that (descriptive) verisimilitude will be an improper measure because any explication will violate the conjunction-principle. In fact, it is not difficult to check that our Definition 2 guarantees the property for two descriptive statements satisfying the conditions.

The importance of his example of (in our sense, theoretical) hypotheses about two different medical treatments lies in the fact that it illustrates perfectly that a lot of scientific practice cannot be understood if one thinks that scientists are aiming at descriptive truth.

Moreover, the example is a convincing illustration of the claim that a measure of merit for such theoretical hypotheses should satisfy the conjunction-principle.

Now it is not difficult to show, as Cohen suggests, that legisimilitude, at least in our explication of it, satisfies the principle.

Fortunately, it is also easy to show that our stronger merit measure retains the property, i.e. if A and B are logically independent, and if B is closer to the theoretical truth than A, then B is closer to the theoretical truth than A & B.

NOTE

* I would like to thank Henk Zandvoort for a number of stimulating discussions about some issues in the article and for his detailed comments in the final stage. I am also grateful to J. van Benthem, R. Cooke, D. Pearce, D. Miller, I. Niiniluoto, G. Oddie and K. Popper for their comments on a note of June 1981 and the last four also for drawing my attention to Miller's article of 1978. Finally, I thank the referee for redrawing my attention to Cohen's paper of 1980 and for stimulating me in adding some clarifying digressions. Of course, all responsibility is mine.

REFERENCES

Cohen, L. J.: 1980, 'What has science to do with truth?', *Synthese* **45**, 489–510.

Lakatos, I.: 1970 'Falsification and the methodology of scientific research programmes', in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*, Cambridge UP, London, pp. 91–196.

Miller, D.: 1974, 'Popper's qualitative theory of verisimilitude', *The British Journal for the Philosophy of Science* **25**, 166–177.

Miller, D.: 1978, 'On distance from the truth as a true distance', in J. Hintikka, I. Niiniluoto, and E. Saarinen (eds.), *Essays on Mathematical and Philosophical Logic*, Reidel, Dordrecht, pp. 415–435.

Nagel, E.: 1961, *The Structure of Science*, Routledge and Kegan Paul, London.

Niiniluoto, I.: 1978, 'Truthlikeness in first order languages', in J. Hintikka, I. Niiniluoto, and

E. Saarinen (eds.), *Essays on Mathematical and Philosophical Logic*, Reidel, Dordrecht, pp. 437–458.
Oddie, G.: 1981, 'Verisimilitude reviewed', *The British Journal for the Philosophy of Science* **32**, 237–265.
Popper, K. R.: 1959, *The Logic of Scientific Discovery*, Hutchinson, London.
Popper, K. R.: 1963, *Conjectures and Refutations*, Routledge and Kegan Paul, London.
Stegmüller, W.: 1979, *The Structuralist View of Theories*, Springer, Berlin.
Tichý, P.: 1974, 'On Popper's definitions of verisimilitude', *The British Journal for the Philosophy of Science* **25**, 155–160.
Tichý, P.: 1978, 'Verisimilitude revisited', *Synthese* **38**, 175–196.

Manuscript submitted 25 May 1982
Final version received 25 August 1982